Available on CMS information server

IEKP-KA/98-21
CMS NOTE 1999/001

The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

21 October 1998

Possibilities of $t\bar{t}H^0$ Event Reconstruction

V. Drollinger, T. Müller

IEKP, Karlsruhe University, Germany

R. Kinnunen

Helsinki Institute of Physics, Helsinki, Finland

Abstract

A light Standard Model Higgs boson can be produced in association with $t\bar{t}$ in the process $pp \rightarrow t\bar{t}H^0$ and detected in the final state $b\bar{b}b\bar{b}q\bar{q}l\nu$. The cross section of this channel is not large, but there are only few sources of Standard Model background with similar event structures. Here we investigate how the correct b jets can be selected in view of H^0 mass reconstruction. We discuss three methods of event reconstruction, a background suppression method for $t\bar{t}$+$jets$ background events, and the effects of jet energy resolution smearing.

1 Physics Processes

In the Standard Model, the H^0 can be produced together with a $t\bar{t}$ pair. The main production mechanism is shown in figure 1a: two gluons split into a $t\bar{t}$ pair, then two of the top quarks annihilate to H^0. With a mass of 100 GeV/c^2 the Higgs boson decays to $b\bar{b}$ with a branching ratio of about 85%. The two associated top quarks decay almost exclusively to a W boson and a bottom quark. In the case we are interested in, one W boson decays into two light quarks, with a branching ratio of 67%, and the second W boson is required to decay into a lepton[1)] (electron or muon) and a neutrino, with a branching ratio of 22%. The final state thus consists of four b jets, two non b jets, one

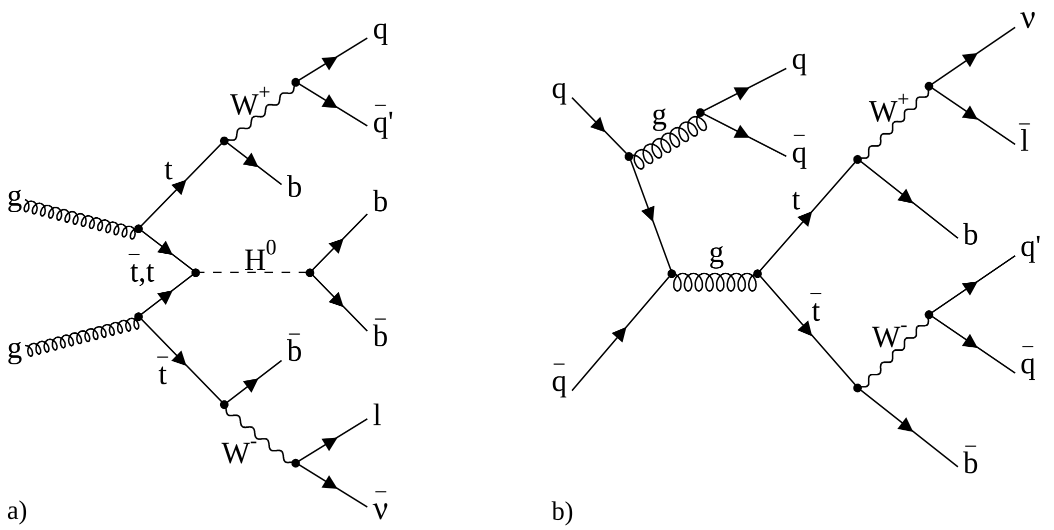


Figure 1: *a) Illustration of Higgs production in association with $t\bar{t}$, with H^0 decaying to $b\bar{b}$. One of the W bosons is required to decay leptonically, the other hadronically. b) $t\bar{t}$ with gluon splitting a possible event of $t\bar{t}$+$jets$ background.*

lepton and missing transverse energy. For an integrated luminosity of $L_{int} = 3 \cdot 10^4 pb^{-1}$ as expected for a "low luminosity run" of the LHC and a H^0 mass of 100 GeV/c^2, about 8400 events with this signature are expected. All events are generated with PYTHIA [1]. For the signal events MSUB 121 and 122 are used[2)].

The main source of background is $t\bar{t}$+$jets$ events, which we analyse here (figure 1b). These events contain four real b jets in case of $t\bar{t}$ with gluon splitting $g \rightarrow b\bar{b}$. Only about 5% of background events from $t\bar{t}$ with gluon splitting have four genuine b jets. In all other cases there are only two b jets. A large background arises from "wrong" reconstruction of jets and the "tagging" of non b jets. For $t\bar{t}$ production MSEL 6 is used.

The CMS detector is modeled by CMSJET [3], a software package for fast detector simulation. The following parameters for the jet searches were used: energy threshold for preselection $E_T > 0.5\ GeV$ per calorimeter cell, precluster energy threshold $E > 5\ GeV$, the jet energy $E_T > 10\ GeV$ and jet eta range $|\eta| < 5$. A jet reconstruction cone of $R_{cone} = 0.4$ is found to be the most appropriate. For soft jets a bigger cone leads in general to a better measurement of the jet energy, but in these events, with a minimum of six jets, a small cone is better to separate the jets from each other. A large cone and harder cuts would make the search for multi-jet events less efficient, because in most cases not all jets can be found. We checked the CMSJET program by comparing reconstructed Z^0 masses in $Z^0 \rightarrow q\bar{q}$ events to the corresponding CMSIM reconstruction as described in [4]. The agreement between the two simulations is very good. Within statistical fluctuation no difference is found for these two-jet events.

1) The efficiency of reconstructing a lepton is assumed to be 90% for muons and electrons.

2) The cross section calculated by PYTHIA is in agreement with the theoretical prediction. The higher order QCD corrections are estimated to increase the cross section by 20% [2].

2 Common Event Selection Features

For all reconstruction methods we use three common selection criteria:

- ⋄lepton trigger
- ⋄requirements on jet number
- ⋄b tagging method

An *isolated* electron or muon with $p_T > 20\ GeV/c$ within $|\eta| < 2.5$ is used as trigger for the event. The lepton is accepted as isolated, when there is no track with $p_T > 2\ GeV/c$ in a cone of radius $R = 0.3\ rad$ around the lepton.

Trigger: 1 *isolated* lepton with $p_T > 20\ GeV/c$ and $|\eta| < 2.5$

For event reconstruction two classes of jets are needed: at least six jets reconstructed with CMSJET as described in the previous chapter. These jets can be within the whole detector η range. At least four of these jets must satisfy the following b tagging conditions: they have to be in the tracker acceptance ($|\eta| < 2.5$) and there must be at least two tracks with $p_T > 1\ GeV/c$ inside the jet reconstruction cone for the impact parameter calculation.

Jets:	min. 6 jets: $E_T > 10\ GeV$ and $|\eta| < 5$
	min. 4 jets: $|\eta| < 2.5$ and min. 2 tracks: $p_T > 1\ GeV/c$

The impact parameter method is used for b tagging as described in [5]: to be tagged as a b jet, two tracks with $p_T > 1\ GeV/c$, impact parameter $ip < 1\ mm$ and a significance of the impact parameter measurement $\sigma(ip) > X$[3)] are required.

b tagging:	min. 2 tracks with	$p_T > 1\ GeV/c$
		$ip < 1\ mm$
		$\sigma(ip) > X$

3 Reconstruction Method I

In events which satisfy the trigger and jet number requirement the W boson decaying hadronically ($W^{\pm}_{hadr.}$) is reconstructed. Only jets with $\sigma(ip) < 2$ are considered[4)] for the calculation of all possible two-jet invariant mass combinations. The mass closest to $70\ GeV/c^2$, the most probable reconstructed W mass value $m_{rec.}$ (no jet energy scaling factors are implemented), is taken as the best reconstructed W mass and the corresponding jets are labelled as jets of the $W^{\pm}_{hadr.}$ decay. This mass has to be within the mass range $40\ GeV/c^2 < m_W < 100\ GeV/c^2$.
In the next step the b jet candidates are selected: the tracks in each jet are sorted according to $\sigma(ip)$. Then the jets are sorted according to the tracks with second highest $\sigma(ip)$[5)]. The four jets with the highest $\sigma(ip)$ are taken as the b jets. At this point all jets for the reconstruction are selected. The possible remaining jets are neglected. For further reconstruction, all four b jets need to have $\sigma(ip) > 2$.
Now everything is prepared for the reconstruction of the first top quark ($t_{hadr.}$). From the two jets assigned to $W^{\pm}_{hadr.}$ and from the four selected b jets four combinations of invariant top masses are calculated. The one with the mass closest to $160\ GeV/c^2$ ($m_{rec.}$ of the top quark) is selected and the corresponding b jet is marked as jet from $t_{hadr.}$, so that there are only three b candidates left for the reconstruction of the rest of the event. The mass window for $t_{hadr.}$ is taken to be $120\ GeV/c^2 < m_{inv.} < 190\ GeV/c^2$.
By using the lepton and the missing transverse energy, the second W boson ($W^{\pm}_{lept.}$) is reconstructed. In this case only the transverse mass of $W^{\pm}_{lept.}$ can be calculated. The event is accepted, if the transverse mass satisfies the condition $30\ GeV/c^2 < m_T < 100\ GeV/c^2$.
The three candidates of transverse mass of the second top quark $t_{lept.}$ are reconstructed using the missing transverse energy, the lepton and one of the three remaining b jets. The transverse mass closest to $130\ GeV/c^2$ of the three

3) X stands for a variable cut, which is different in each case. A cut of $\sigma(ip) > 3$ would be usual for b tagging of a single jet.

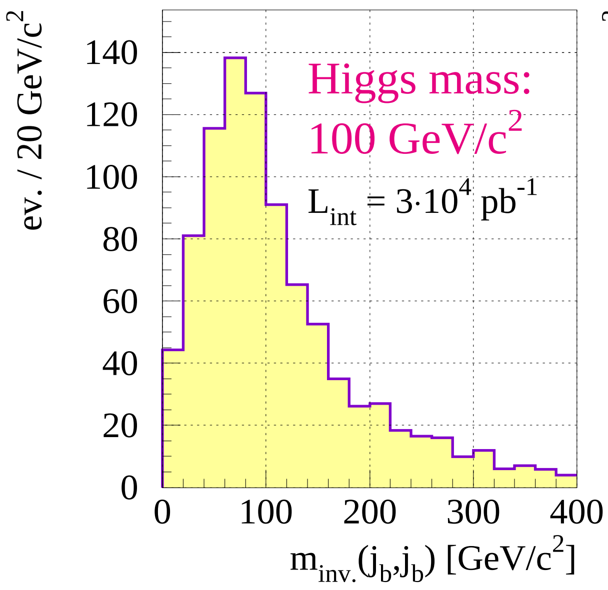
4) This cut is used, because no b jet candidates should be used for the W boson reconstruction.

5) This means, that the jet would be tagged with a cut lower as this value, because only two tracks are required in each jet. So the jets are most b like, when $\sigma(ip)$ of their track with the second highest $\sigma(ip)$ has a high value.

combinations is retained as the reconstructed top mass, if it is between 90 GeV/c^2 and 170 GeV/c^2, otherwise the event is not accepted. The b jet which is used for the $t_{lept.}$ reconstruction is also marked as "already used".
Now there are only two b candidates left. These are supposed to be the jets coming from the H^0 decay. In this last step the invariant mass of the last two b candidates is calculated. Additional cuts on these last jets are of no help to get a better signal or to improve the significance.
The following table contains information about each reconstruction step and its efficiencies for a H^0 signal with a mass of 100 GeV/c^2 and for $t\bar{t}$+$jets$ background events:

Method I: Acceptances for $L_{int} = 3 \cdot 10^4 pb^{-1}$		
Reconstruction Steps	Signal: H^0 (100 GeV/c^2)	Background: $t\bar{t}$+$jets$
Triggered Events with min. 6 and 4 b Jets	4400	1670000
Reconstruction of $W^{\pm}_{hadr.}$	85%	90%
Selection of four b Candidates	48%	17%
Reconstruction of $t_{hadr.}$	84%	80%
Reconstruction of $W^{\pm}_{lept.}$	69%	65%
Reconstruction of $t_{lept.}$	89%	89%
Number of selected Events	930	118000
Reconstruction Efficiency	21%	7.1%

The previous table shows that the reconstruction does help to suppress background, although not very effectively, but is necessary to select the final two b candidates for H^0 mass reconstruction. *The Higgs signal significance as obtained for this method is* $S/\sqrt{B} = 2.7$ *for* $L_{int} = 3 \cdot 10^4 pb^{-1}$. In figure 2 the $b - b$ invariant mass distributions of the signal and $t\bar{t}$+$jets$ background are shown. A "peaklike behaviour" of the signal distribution is not very pronounced, the background invariant mass distribution is just somewhat wider.



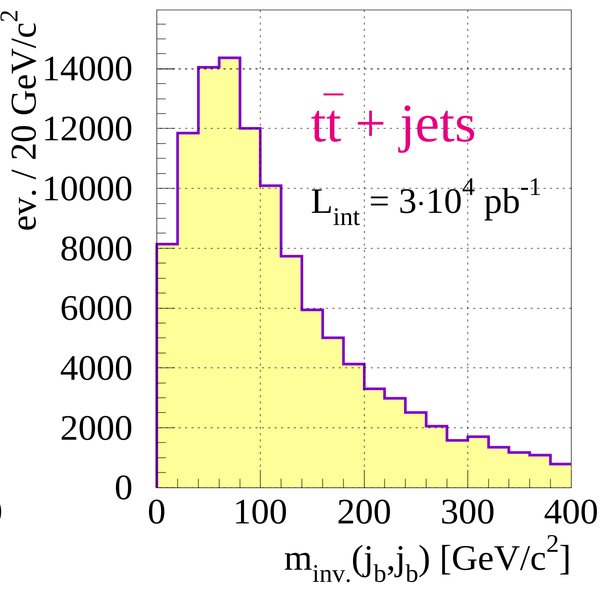


Figure 2: *Invariant mass distributions reconstructed from the two finally selected b-jets for H^0 signal and $t\bar{t}$+$jets$ background, with method I.*

4 Reconstruction Method II

Method II uses a different b jet selection and top reconstruction procedure. The two W bosons $W^{\pm}_{lept.}$ and $W^{\pm}_{hadr.}$ are reconstructed as in method I. Also the selection of the four b candidates with impact parameter tagging is the

same as in method I. But then the four b jets are ordered according to E_T. The two jets with the highest jet E_T are now identified with the b jets from top decay. The two b candidates with the lower E_T are treated as the jets coming from the H^0. The reason for this selection criterion is the lower H^0 mass compared to the top mass. So the bottom quarks from top decays should have on average a higher transverse energy. To improve the quality of the final signal, we applied a harder b tagging cut $\sigma(ip) > 3$ and a ΔR cut $\Delta R > 0.8\ rad$ between the jets from the H^0 against gluon splitting processes. Now the invariant mass of the two selected b-jets can be calculated.
The reconstructed $t_{hadr.}$ and $t_{lept.}$ masses are calculated from the decay products of the W bosons and from the two highest E_T b candidates. There are two possibilities to reconstruct both top quarks: for $t_{hadr.}$ and for $t_{lept.}$ the combination nearest to $m_{inv.}(hadr.) = 160\ GeV/c^2$ and $m_T(lept.) = 130\ GeV/c^2$ is kept. Events now have to pass the following two mass cuts: for $t_{hadr.}$ $120\ GeV/c^2 < m_{inv.} < 200\ GeV/c^2$ and for $t_{lept.}$ $100\ GeV/c^2 < m_T < 180\ GeV/c^2$.

Method II: Acceptances for $L_{int} = 3 \cdot 10^4 pb^{-1}$		
Reconstruction Steps	Signal: H^0 ($100\ GeV/c^2$)	Background: $t\bar{t}+jets$
Triggered Events with min. 6 and 4 b Jets	4400	1670000
Reconstruction of $W^{\pm}_{lept.}$	70%	65%
Reconstruction of $W^{\pm}_{hadr.}$	85%	90%
Selection of four b Candidates	48%	17%
Second b tagging of H^0 Candidates	58%	43%
$\Delta R_{jet,jet}$ Cut	79%	72%
Reconstruction of $t_{hadr.}$ and $t_{lept.}$	38%	33%
Number of selected Events	217	16500
Reconstruction Efficiency	4.9%	0.98%

Figure 3 shows the invariant mass distributions of the H^0 signal and of $t\bar{t}+jets$ background. In spite of the detailed selection and not very efficient reconstruction, the distributions display no big difference. *The statistical significance of method II is $S/\sqrt{B} = 1.7$ for $L_{int} = 3 \cdot 10^4 pb^{-1}$*. For a better significance a sharper signal peak would be needed, which would allow cutting in the invariant mass distribution.

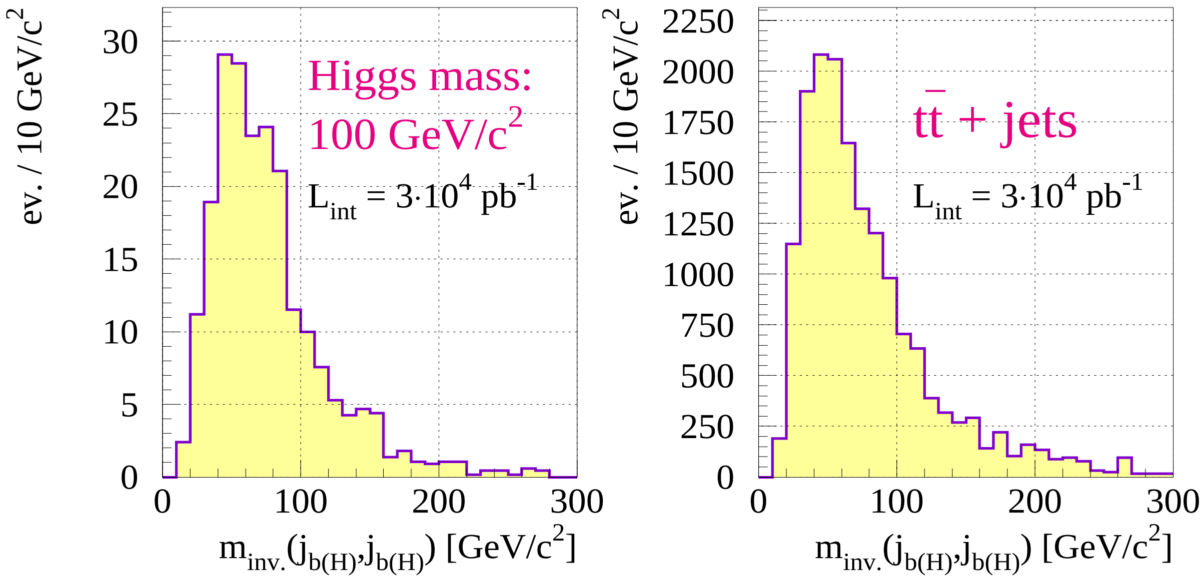


Figure 3: *Invariant mass distributions of the H^0 signal on the left side and $t\bar{t}+jets$ background on the right reconstructed with method II.*

5 Reconstruction Method III

In contrast to the previous methods, neither W bosons nor the top quarks are reconstructed in method III. Here we calculate simply all combinations of distance between two jets $\Delta R_{jet,jet}$ and the invariant two-jet mass $m_{inv.}$ from the pairs of selected jets. The triggering and jet selection is the same as before.

As a first step, we select the jets to be used for the calculation of all combinations. For impact parameter tagging $\sigma(ip) > 3$ is required. After that a certain number of jets is left. The event is accepted only, if there are at least three jets selected for the calculations of the next step.

In this part of the reconstruction $\Delta R_{jet,jet}$ and $m_{inv.}$ of each pair of selected jets are calculated for all possible ($N_{comb.} = 0.5 \times (n_{jets} - 1) \times n_{jets}$) combinations. Then the combinations of interest are selected with the $\Delta R_{jet,jet}$ condition: $0.6\, rad < \Delta R_{jet,jet} < 2.8\, rad.$

Then the number of combinations left is counted. If there is at least one combination left, we calculate a weight ($weight = N_{left}^{-1}$). Now each combination of an event can be counted or histogrammed with this weight. With this method one does not lose any event through the $\Delta R_{jet,jet}$ requirement, when there is left at least one combination. This improves significantly the efficiency.

Method III: Acceptances for $L_{int} = 3 \cdot 10^4 pb^{-1}$		
Reconstruction Steps	Signal: H^0 (100 GeV/c^2)	Background: $t\bar{t}$+$jets$
Triggered Events with min. 6 and 4 b Jets	4400	1670000
Cut on Number of selected Jets	55%	22%
$\Delta R_{jet,jet}$ Cut of all Combinations	97%	91%
Number of selected Events	2340	342000
Reconstruction Efficiency	53%	20%

In figure 4 one can see the results of this method: the invariant mass distributions for signal and background. In the signal distribution it is possible to see a hint of a peak, but this is not yet significant enough to allow cuts in the invariant mass. *The significance of this method is* $S/\sqrt{B} = 4.0$ *for* $L_{int} = 3 \cdot 10^4 pb^{-1}$. The reason for this more promising result is the high efficiency of the event selection.

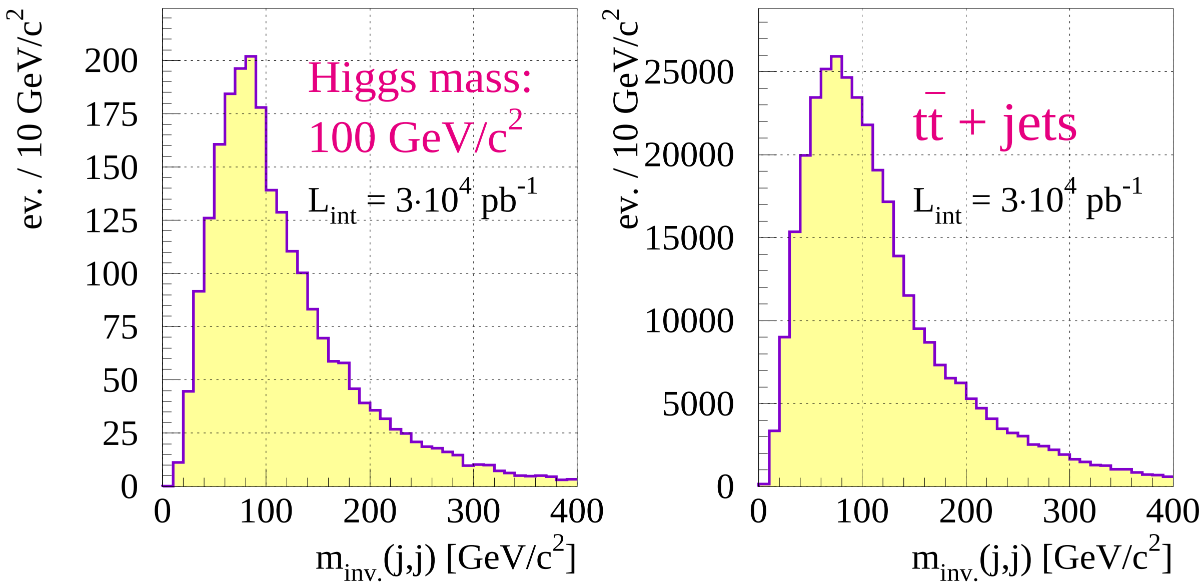


Figure 4: *The invariant mass distributions of selected jet pairs, reconstructed with method III, for H^0 signal (left) and $t\bar{t}$+$jets$ background (right).*

6 Jet Smearing Effects

None of the methods presented above leads yet to a satisfying result as far as the visibility of a signal mass peak at or near the H^0 mass is concerned. Now we try to investigate the reasons for this behaviour. There are two main possibilities that could be responsible for the disappearance of the peak of the H^0 signal:

◇The measurement of the jet energy is not precise enough.
◇The event reconstruction is not working for multi-jet events.

To investigate the first point, the momenta of the partons producing a jet are smeared with a gaussian. The loss of energy is taken into account as a simple factor. An energy loss factor $E_{loss} = 80\%$ and $\sigma_{Gauss} = 1.8$ ($\Leftrightarrow \Delta E/E = 180\%/\sqrt{E}$) leads to a result comparable to the mass resolution of $Z^0 \to q\bar{q}$ events as obtained with CMSJET [3] or with CMSIM [4]. The rest of the detector simulation (lepton, b tagging) is not changed.
The second point is investigated with a CMSJET run without any energy smearing switched on. In this case only the spatial resolution effect included by the size of different calorimeter cells and the jet reconstruction algorithm can smear the jet energy.

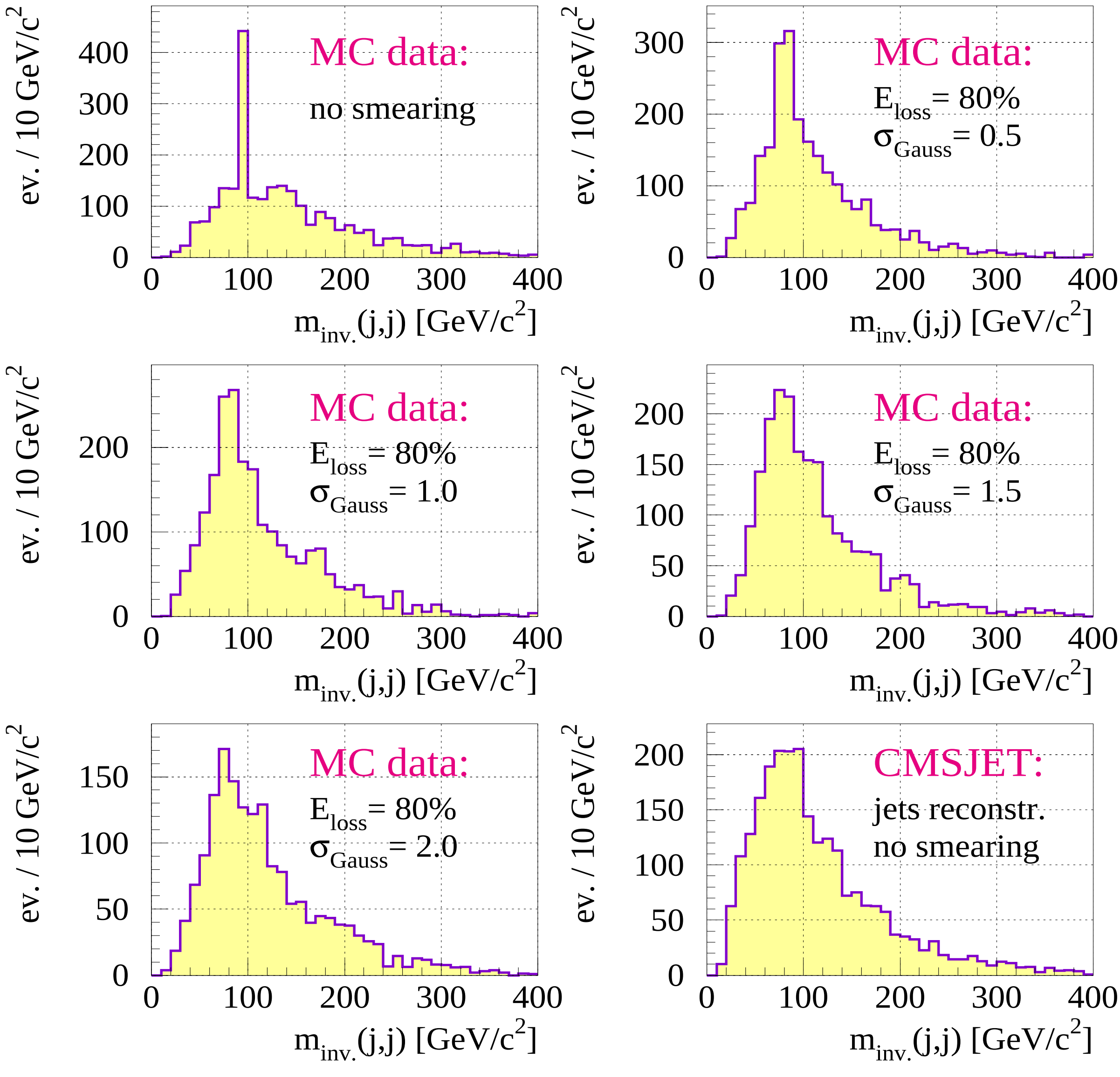


Figure 5: *Invariant mass distributions for signal events have been reconstructed with method III using smeared partons of the Monte Carlo events, and reconstructed jets of CMSJET without any energy smearing in the last histogram. With increasing σ_{Gauss} the peak at or near the input H^0 mass becomes wider. With increasing σ_{Gauss} it leads to poorer visibility of the signal peak. The last figure illustrates the difficulties of jet reconstruction in multi-jet events.*

In figure 5 the different smearing effects are illustrated for an H^0 signal at $100\ GeV/c^2$ (method III): the plot with the unsmeared Monte Carlo information shows that the jet selection with b tagging works quite well. There is a clear signal peak above the combinatorial background. The other (not shown) reconstruction methods also lead to a good result with these assumptions.
With increasing σ_{Gauss} of the parton energy smearing the signal peak is getting wider and lower and nearly disappears at $\sigma_{Gauss} = 2.0$ (the left plot at the bottom). E_{loss} has a minor effect on the mass resolution. It only shifts the peak to the left.
The second test with "unsmeared" CMSJET shows that the algorithm is not working very well for multi-jet events as can already be seen from figure 4. The smearing of the H^0 mass peak is as large as that from the jet energy measurement smearing.

7 Summary and Conclusions

We have presented three methods to reconstruct $t\bar{t}H^0$ events. In method I and method II the two W bosons and the two top quarks are reconstructed from the lepton, the missing transverse energy and jets selected either by impact parameter tagging or by hardness of the jet transverse momentum. No clear peak is visible in the invariant mass distribution for H^0. Because these methods are not very efficient, the significances are not high.
In the simpler method III the individual objects ($W^{\pm}_{lept.}$, $W^{\pm}_{hadr.}$, $t_{hadr.}$ and $t_{lept.}$) are not reconstructed. All possible combinations of selected b jet candidates are used for H^0 reconstruction. The significance is better here due to the higher reconstruction efficiency, nevertheless no clear signal peak is visible yet.

Comparison of all Methods			
Reconstruction Method	Rec. Efficiency: $t\bar{t}H^0$	Rec. Efficiency: $t\bar{t}+jets$	Significance $S/\sqrt{B}$
Methode I	21%	7.1%	2.7
Methode II	4.9%	0.98%	1.7
Methode III	53%	20%	4.0

For all methods investigated the background suppression is not good enough - only b tagging is really useful to suppress the large background. Although the irreducible background from $b\bar{b}b\bar{b}q\bar{q}l\nu$ events is small, a lot of background is picked up from $t\bar{t}+jets$ events with false b tags and a bad jet reconstruction. Also the event reconstruction efficiency suffers from poor jet resolution.

We find two main reasons for the poor H^0 mass resolution: the jet energy resolution and the difficulty of jet reconstruction in multi-jet events. A better energy resolution of the calorimeters (HCAL and VFCAL) could be helpful to get better results. A better jet reconstruction algorithm is necessary to reconstruct the $t\bar{t}H^0$ channel and multi-jet physics in general.

The study of b tagging performance as well as of methods of jet reconstruction are in progress with a detailed simulation using CMSIM. This with improved overall event reconstruction algorithms using likelihood methods for example will hopefully lead to new possibilities of $t\bar{t}H^0$ event reconstruction.

We thank D. Denegri for his many suggestions and interest in this work!

References

[1] **CERN-TH.7112/93**,T. Sjostrand “PYTHIA 5.7 and JETSET 7.4 Physics and manual”

[2] **CERN-TH Workshop**, M.Spira “Production Properties of Higgs and SUSY Particles” (http://wwwth.cern.ch/lhcworkshop98/ps/program.html)

[3] **CMS TN/94-180**, S. Abdullin, A. Khanov and N. Stepanov, “CMSJET”

[4] **CMS NOTE-1998/026**, D.Green, K.Maeshima, J.Wormersley and W.Wu “Dijet Mass Resolution of the CMS Calorimeter”

[5] **CMS TN/96-045**, R. Kinnunen and D. Denegri “B-tagging with Impact Parameter in the CMS Tracker”

[6] **CMS Collaboration, Technical Proposal,CERN/LHCC 94-38, LHCC/P1, 15 December 1994**

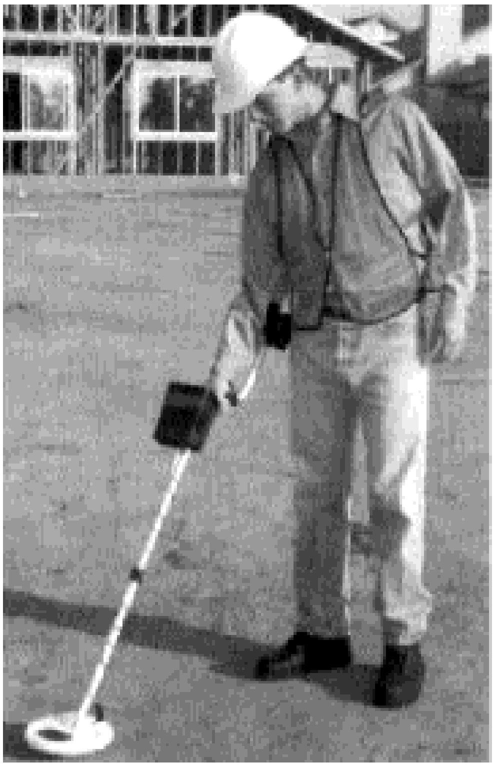


Figure 6: *Higgs Searching with VCMS*[6)]

[6)] Very Compact Muon Solenoid